



1600

RAW SEQUENCE LISTING

DATE: 02/19/2003

PATENT APPLICATION: US/09/720,529A

TIME: 13:51:24

Input Set : A:\BB-1118-A Corrected Seq List - EMO.txt

Output Set: N:\CRF4\02192003\I720529A.raw

3 <110> APPLICANT: Cahoon, Rebecca E.
 4 Vollmer, Steven J.
 6 <120> TITLE OF INVENTION: Chromatin Associated Proteins
 8 <130> FILE REFERENCE: BB-1118-A
 10 <140> CURRENT APPLICATION NUMBER: US 09/720,529A
 11 <141> CURRENT FILING DATE: 2000-12-20
 13 <150> PRIOR APPLICATION NUMBER: US 60/092,841
 14 <151> PRIOR FILING DATE: 1998-07-14
 16 <150> PRIOR APPLICATION NUMBER: PCT/US99/15807
 17 <151> PRIOR FILING DATE: 1999-07-13
 19 <160> NUMBER OF SEQ ID NOS: 8
 21 <170> SOFTWARE: Microsoft Office 97
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 1990
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Oryza sativa
 28 <400> SEQUENCE: 1

29	gctacttcta	cgacgcggag	gtggggaact	actactacgg	gcagggggcac	ccgatgaagc	60
30	cgcaccgcat	ccggatgacc	cacgcgctgc	tcgcccacta	cggcctcctc	gaccagatgc	120
31	agggtgctcaa	gccccacccg	gcgcgcgacc	gcgacctctg	ccgcttccac	gccgacgact	180
32	acgtcgctt	cctccgctcc	gtcacgcggg	agaccagca	ggaccagatc	cgggcgctca	240
33	agcgcttcaa	cgtcggcgag	gactgccccg	tcttcgacgg	cctctacagc	ttctgccaga	300
34	cctacgccgg	gggatccgct	ggcggcgccg	tcaagctcaa	ccacggccac	gacatcgcca	360
35	tcaactgggc	cggcggcctc	caccacgcca	agaagtgcga	ggcctcgga	ttctgctacg	420
36	tcaacgacat	cgctcctcgcc	atcctcgagc	tcctcaaata	ccaccagcgt	gttctctatg	480
37	tggatatcga	tatccaccat	ggggatggtg	tggaggaggc	gttctacacg	acggacaggg	540
38	tgatgacggt	ctcggttccac	aagtttgggg	attatttccc	ggggaccggg	gacattcgcg	600
39	atattgggca	ctcaaagggg	aagtattact	ctctgaatgt	cccgttgga	gacggtatcg	660
40	acgacgagag	ctaccagtcg	ttgttcaagc	cgatcatggg	gaaggtgatg	gaggttttct	720
41	gccctggcgc	ggtggtgctc	cagtgcggtg	cggactctct	gtcgggtgat	aggttgggtt	780
42	gcttcaacct	gtcaatcagg	ggccacgcgg	aatgcgtgag	attcatgagg	tccttcaatg	840
43	tcccgcgtgt	gctgcttggt	ggtggtgggt	ataccataag	aaatgttgcg	cgggtgtggt	900
44	gctatgagac	aggagttgca	cttggtcatg	agctcactga	caagatgcct	ccaaatgagt	960
45	attttgagta	ctttggtcca	gattatacac	ttcatgttgc	accaagtaac	atggagaaca	1020
46	aaaacacacg	ccagcagttg	gatgatataa	gatcaagact	tcttgataat	ctttcaaaac	1080
47	ttcgacatgc	tcctagcgct	caatttcaag	agcgaccccc	tgaggctgag	ctacctgagc	1140
48	aagatgaaga	ccaagaggat	cctgatgaaa	ggcaccatgc	tgattctgat	gtggaaatgg	1200
49	atgatgtcaa	acctttggat	gactcaggaa	ggaggagcag	tattcagaat	gtgagagtta	1260
50	agagagagtc	tgctgaaaca	gatgccgcag	atcaggatgg	taatagggtc	gctgcagaga	1320
51	acaccaaggg	cacagaacct	gcggctgatg	gagttggttc	ctcgaaacaa	actgttccta	1380
52	ccgatgcaag	tgcatggcc	atagacgaac	caggctccct	gaaagtcgag	ccagataact	1440
53	caaacaaatt	gcaagatcaa	ccatcggtgc	accagaagac	ataatagttc	tctctacctt	1500
54	aaaacttagt	aactgatgcc	atctatcatc	cattgattat	attggagaaa	ctcccaactt	1560

ENTERED

RAW SEQUENCE LISTING

DATE: 02/19/2003

PATENT APPLICATION: US/09/720,529A

TIME: 13:51:24

Input Set : A:\BB-1118-A Corrected Seq List - EMO.txt

Output Set: N:\CRF4\02192003\I720529A.raw

```

55 tgaagcagag agttcatgcc ataccaaaag ttatatacca aatttcgaat ggtatgtaca 1620
56 cctttcgaac tgggtggtgtt ttgtgcaata catttatgcc aggctgacta ttatgtggta 1680
57 tctattatta gctttagttt aaccctgtct gctgtcgagc aatcgggata gtcgtgcaat 1740
58 atattctgga tctatcaagc aatgtgagac ggatgtcaaa ccattgggtg tgacttcagc 1800
59 aatgtatgta tatgtaagta tagggacagg cagcaggcgt tactttggtg gaagctacaa 1860
60 gctttgtctc tcttctcatc cctaatacct acgtggggtg cgtctcgttg ttgtttaggt 1920
61 attactgcat tcttaaaagt gctcatttag ggtgaaattc taacatcttc taaaaaaaaa 1980
62 aaaaaaaaaa 1990

```

64 <210> SEQ ID NO: 2

65 <211> LENGTH: 493

66 <212> TYPE: PRT

67 <213> ORGANISM: Oryza sativa

69 <400> SEQUENCE: 2

```

70 Tyr Phe Tyr Asp Ala Glu Val Gly Asn Tyr Tyr Tyr Gly Gln Gly His
71   1           5           10           15
73 Pro Met Lys Pro His Arg Ile Arg Met Thr His Ala Leu Leu Ala His
74           20           25           30
76 Tyr Gly Leu Leu Asp Gln Met Gln Val Leu Lys Pro His Pro Ala Arg
77           35           40           45
79 Asp Arg Asp Leu Cys Arg Phe His Ala Asp Asp Tyr Val Ala Phe Leu
80           50           55           60
82 Arg Ser Val Thr Pro Glu Thr Gln Gln Asp Gln Ile Arg Ala Leu Lys
83   65           70           75           80
85 Arg Phe Asn Val Gly Glu Asp Cys Pro Val Phe Asp Gly Leu Tyr Ser
86           85           90           95
88 Phe Cys Gln Thr Tyr Ala Gly Gly Ser Val Gly Gly Ala Val Lys Leu
89           100          105          110
91 Asn His Gly His Asp Ile Ala Ile Asn Trp Ala Gly Gly Leu His His
92           115          120          125
94 Ala Lys Lys Cys Glu Ala Ser Gly Phe Cys Tyr Val Asn Asp Ile Val
95           130          135          140
97 Leu Ala Ile Leu Glu Leu Leu Lys Tyr His Gln Arg Val Leu Tyr Val
98   145          150          155          160
100 Asp Ile Asp Ile His His Gly Asp Gly Val Glu Glu Ala Phe Tyr Thr
101           165          170          175
103 Thr Asp Arg Val Met Thr Val Ser Phe His Lys Phe Gly Asp Tyr Phe
104           180          185          190
106 Pro Gly Thr Gly Asp Ile Arg Asp Ile Gly His Ser Lys Gly Lys Tyr
107           195          200          205
109 Tyr Ser Leu Asn Val Pro Leu Asp Asp Gly Ile Asp Asp Glu Ser Tyr
110          210          215          220
112 Gln Ser Leu Phe Lys Pro Ile Met Gly Lys Val Met Glu Val Phe Arg
113   225          230          235          240
115 Pro Gly Ala Val Val Leu Gln Cys Gly Ala Asp Ser Leu Ser Gly Asp
116           245          250          255
118 Arg Leu Gly Cys Phe Asn Leu Ser Ile Arg Gly His Ala Glu Cys Val
119           260          265          270
121 Arg Phe Met Arg Ser Phe Asn Val Pro Leu Leu Leu Leu Gly Gly Gly
122           275          280          285

```

RAW SEQUENCE LISTING

DATE: 02/19/2003

PATENT APPLICATION: US/09/720,529A

TIME: 13:51:24

Input Set : A:\BB-1118-A Corrected Seq List - EMO.txt

Output Set: N:\CRF4\02192003\I720529A.raw

```

124 Gly Tyr Thr Ile Arg Asn Val Ala Arg Cys Trp Cys Tyr Glu Thr Gly
125      290                      295                      300
127 Val Ala Leu Gly His Glu Leu Thr Asp Lys Met Pro Pro Asn Glu Tyr
128 305                      310                      315                      320
130 Phe Glu Tyr Phe Gly Pro Asp Tyr Thr Leu His Val Ala Pro Ser Asn
131      325                      330                      335
133 Met Glu Asn Lys Asn Thr Arg Gln Gln Leu Asp Asp Ile Arg Ser Arg
134      340                      345                      350
136 Leu Leu Asp Asn Leu Ser Lys Leu Arg His Ala Pro Ser Val Gln Phe
137      355                      360                      365
139 Gln Glu Arg Pro Pro Glu Ala Glu Leu Pro Glu Gln Asp Glu Asp Gln
140      370                      375                      380
142 Glu Asp Pro Asp Glu Arg His His Ala Asp Ser Asp Val Glu Met Asp
143 385                      390                      395                      400
145 Asp Val Lys Pro Leu Asp Asp Ser Gly Arg Arg Ser Ser Ile Gln Asn
146      405                      410                      415
148 Val Arg Val Lys Arg Glu Ser Ala Glu Thr Asp Ala Ala Asp Gln Asp
149      420                      425                      430
151 Gly Asn Arg Val Ala Ala Glu Asn Thr Lys Gly Thr Glu Pro Ala Ala
152      435                      440                      445
154 Asp Gly Val Gly Ser Ser Lys Gln Thr Val Pro Thr Asp Ala Ser Ala
155      450                      455                      460
157 Met Ala Ile Asp Glu Pro Gly Ser Leu Lys Val Glu Pro Asp Asn Ser
158 465                      470                      475                      480
160 Asn Lys Leu Gln Asp Gln Pro Ser Val His Gln Lys Thr
161      485                      490
163 <210> SEQ ID NO: 3
164 <211> LENGTH: 1805
165 <212> TYPE: DNA
166 <213> ORGANISM: Glycine max
168 <400> SEQUENCE: 3
169 ctactgcact ctgacactat cagatatcaa tttcagaggc ggactggagg ctggagctgc 60
170 ttcaagtctt caactatctg aatttctgaa ctcattccatt cgtagcgca stagaacgaa 120
171 aactgagtaa tggaaagtgg agggaaactcc cttccatcag ggtcagatgg tgtgaagaga 180
172 aaggtttcat atttctatga tccagagggt ggaaactatt attatgggca gggacaccca 240
173 atgaaaccac acaggattcg aatgacacat gctcttttag cccactatgg attgcttcaa 300
174 cacatgcagg ttctgaagcc tatggctgct aaagataggg accttgcaa gttccatgct 360
175 gatgattatg tggcctttct gagaggcatc acccctgaaa cgcagcaaga tcaattgaga 420
176 cagctgaaga ggtttaatgt tggcgaagac tgccctgtat ttgatggtct ttactctttc 480
177 tgccagacat atgcaggagg ttctgttggt ggtgctctaa agttgaacca tggagtatgt 540
178 gatattgcaa taaattgggc tgggtgtcta catcatgcaa agaagtgtga ggcttctggg 600
179 ttttgctatg ttaatgacat tgtgtgtggt attttggaac ttctcaaaat acatgagcgt 660
180 gttctgtatg tggacattga tatccacat ggtgatggtg tagaggaggc cttttacacc 720
181 accgataggg tcatgactgt ttcgtttcat aagtttgggg attactttcc tggacacagg 780
182 gatattcgtg atattggata tgctaaaggg aatatattatt cactaaatgt tcccttgga 840
183 gatggaattg atgatgagag ctatcagtc tggtttaagc caataatggg aaagggtatg 900
184 gagattttta ggcccgggtg tgttgattta caatgtggtg ctgactcttt atctggggac 960
185 aggttaggtt gtttcaatct ttccataaaa ggtcatgcag agtgtgtcag atatatgaga 1020
186 tcttttaatg ttccccttct attgctcggg ggagggtggt atacaataag aaatgtggca 1080

```

RAW SEQUENCE LISTING

DATE: 02/19/2003

PATENT APPLICATION: US/09/720,529A

TIME: 13:51:24

Input Set : A:\BB-1118-A Corrected Seq List - EMO.txt.

Output Set: N:\CRF4\02192003\I720529A.raw

```

187 cgttgttggt gtttcgagac tagcgttgct cttgggattg aactagatga taagatgcct 1140
188 caacatgaat attatgaata ttttggtcct gactatactc ttcatgttgc tccaagtaac 1200
189 atggaaaaca agaactcccg acaattattg gatgaaataa gagcaaaact tcttgataat 1260
190 ttatctaggc ttcaacatgc accaagtgtc ccattccagg aacggccacc tgatgcagag 1320
191 cttctagaga gagatgaaga tcaagatgat agagatgaaa gatgggatcc cgattctgac 1380
192 agrgagggtg gcgatgacag caatcctgtt cgcagaaggg tgaaaagtga atgcgttgat 1440
193 gctgaggata aagatacagt gtcgggggtg gactcaatgg cagtggatga accatgcac 1500
194 aaagaggagc aggataattht aaaagagctt tctgatacaca ggccaagatg aagcaataat 1560
195 aagcattact gatcaacctt ctctctgact agtgctctgtc gacctgtaaa ttatagtttc 1620
196 ctctaaagc agtctggcat gcattcatct gacgtctgta gtgtttcaaa tttttgcttt 1680
197 atctggaaac tgaagagata tggtgcaagc ttgccttggc ttttgatgtt tcatattact 1740
198 gcaagatgaa tgtagtagtt attttttctg taaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800
199 aaaaaa 1805
201 <210> SEQ ID NO: 4
202 <211> LENGTH: 473
203 <212> TYPE: PRT
204 <213> ORGANISM: Glycine max
206 <400> SEQUENCE: 4
207 Met Glu Ser Gly Gly Asn Ser Leu Pro Ser Gly Ser Asp Gly Val Lys
208 1 5 10 15
210 Arg Lys Val Ser Tyr Phe Tyr Asp Pro Glu Val Gly Asn Tyr Tyr Tyr
211 20 25 30
213 Gly Gln Gly His Pro Met Lys Pro His Arg Ile Arg Met Thr His Ala
214 35 40 45
216 Leu Leu Ala His Tyr Gly Leu Leu Gln His Met Gln Val Leu Lys Pro
217 50 55 60
219 Met Ala Ala Lys Asp Arg Asp Leu Cys Lys Phe His Ala Asp Asp Tyr
220 65 70 75 80
222 Val Ala Phe Leu Arg Gly Ile Thr Pro Glu Thr Gln Gln Asp Gln Leu
223 85 90 95
225 Arg Gln Leu Lys Arg Phe Asn Val Gly Glu Asp Cys Pro Val Phe Asp
226 100 105 110
228 Gly Leu Tyr Ser Phe Cys Gln Thr Tyr Ala Gly Gly Ser Val Gly Gly
229 115 120 125
231 Ala Leu Lys Leu Asn His Gly Val Cys Asp Ile Ala Ile Asn Trp Ala
232 130 135 140
234 Gly Gly Leu His His Ala Lys Lys Cys Glu Ala Ser Gly Phe Cys Tyr
235 145 150 155 160
237 Val Asn Asp Ile Val Leu Ala Ile Leu Glu Leu Leu Lys Ile His Glu
238 165 170 175
241 Arg Val Leu Tyr Val Asp Ile Asp Ile His His Gly Asp Gly Val Glu
242 180 185 190
244 Glu Ala Phe Tyr Thr Thr Asp Arg Val Met Thr Val Ser Phe His Lys
245 195 200 205
247 Phe Gly Asp Tyr Phe Pro Gly Thr Gly Asp Ile Arg Asp Ile Gly Tyr
248 210 215 220
250 Ala Lys Gly Lys Tyr Tyr Ser Leu Asn Val Pro Leu Asp Asp Gly Ile
251 225 230 235 240
253 Asp Asp Glu Ser Tyr Gln Ser Leu Phe Lys Pro Ile Met Gly Lys Val

```

RAW SEQUENCE LISTING

DATE: 02/19/2003

PATENT APPLICATION: US/09/720,529A

TIME: 13:51:24

Input Set : A:\BB-1118-A Corrected Seq List - EMO.txt

Output Set: N:\CRF4\02192003\I720529A.raw

```

254          245          250          255
256 Met Glu Ile Phe Arg Pro Gly Ala Val Val Leu Gln Cys Gly Ala Asp
257          260          265          270
259 Ser Leu Ser Gly Asp Arg Leu Gly Cys Phe Asn Leu Ser Ile Lys Gly
260          275          280          285
262 His Ala Glu Cys Val Arg Tyr Met Arg Ser Phe Asn Val Pro Leu Leu
263          290          295          300
265 Leu Leu Gly Gly Gly Gly Tyr Thr Ile Arg Asn Val Ala Arg Cys Trp
266 305          310          315          320
268 Cys Phe Glu Thr Ser Val Ala Leu Gly Ile Glu Leu Asp Asp Lys Met
269          325          330          335
271 Pro Gln His Glu Tyr Tyr Glu Tyr Phe Gly Pro Asp Tyr Thr Leu His
272          340          345          350
274 Val Ala Pro Ser Asn Met Glu Asn Lys Asn Ser Arg Gln Leu Leu Asp
275          355          360          365
277 Glu Ile Arg Ala Lys Leu Leu Asp Asn Leu Ser Arg Leu Gln His Ala
278          370          375          380
280 Pro Ser Val Pro Phe Gln Glu Arg Pro Pro Asp Ala Glu Leu Leu Glu
281 385          390          395          400
283 Arg Asp Glu Asp Gln Asp Asp Arg Asp Glu Arg Trp Asp Pro Asp Ser
284          405          410          415
286 Asp Arg Glu Val Gly Asp Asp Ser Asn Pro Val Arg Arg Arg Val Lys
287          420          425          430
289 Ser Glu Cys Val Asp Ala Glu Asp Lys Asp Thr Val Ser Gly Val Asp
290          435          440          445
292 Ser Met Ala Val Asp Glu Pro Cys Ile Lys Glu Glu Gln Asp Asn Leu
293          450          455          460
295 Lys Glu Leu Ser Asp His Arg Pro Arg
296 465          470
298 <210> SEQ ID NO: 5
299 <211> LENGTH: 541
300 <212> TYPE: DNA
301 <213> ORGANISM: Triticum aestivum
303 <220> FEATURE:
304 <221> NAME/KEY: unsure
305 <222> LOCATION: (134)
306 <223> OTHER INFORMATION: n is a, c, g or t
308 <220> FEATURE:
309 <221> NAME/KEY: unsure
310 <222> LOCATION: (387)
311 <223> OTHER INFORMATION: n is a, c, g or t
313 <220> FEATURE:
314 <221> NAME/KEY: unsure
315 <222> LOCATION: (417)
316 <223> OTHER INFORMATION: n is a, c, g or t
318 <220> FEATURE:
319 <221> NAME/KEY: unsure
320 <222> LOCATION: (433)
321 <223> OTHER INFORMATION: n is a, c, g or t

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/720,529A

DATE: 02/19/2003
TIME: 13:51:25

Input Set : A:\BB-1118-A Corrected Seq List - EMO.txt
Output Set: N:\CRF4\02192003\I720529A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 134, 387, 417, 433, 464, 475, 479, 495, 505, 508
Seq#:6; Xaa Pos. 24, 108, 118